**Link to publisher's version:** *https://doi.org/10.21278/IDC.2018.0432*

**Citation:** Duraiswamy V, Campean F, Harris S and Munive-Hernandez E (2018) Development of a methodology for robust evaluation of perceived quality of vehicle body panel gaps. 15th International Design Conference (Proceedings) - Design 2018. 21-24 May 2018. Dubrovnik, Croatia.

INTERNATIONAL DESIGN CONFERENCE - DESIGN 2018
Dubrovnik - Croatia, May 21-24, 2018.

# DEVELOPMENT OF A METHODOLOGY FOR ROBUST EVALUATION OF PERCEIVED QUALITY OF VEHICLE BODY PANEL GAPS

Vidya Duraiswamy, Felician Campean, Steve Harris, Eduardo Munive-Hernandez

## Abstract

This paper presents research into perceived quality of vehicle body split lines / gaps. The survey based methodology combined direct attribute evaluation and choice experiments with multiple test cases, based on static images generated from parametric CAD virtual prototypes of a specific vehicle. The study systematically tested for perception without and with awareness and the Hawthorne bias, showing significant effects. Statistical analysis provided clear evidence of the user preference for smaller gap size, and as such supports target and tolerance setting for the body gap.

## 1. Introduction

Perceived quality is defined in relation to the customers' emotional response to the aesthetic appearance of a product, including fit and finish, appearance of surfaces, craftsmanship and execution of details (Wickman et al, 2014). For automotive products, perceived quality is significantly influenced by manufacturing quality design, including craftsmanship and gap and flush measurements of vehicle body split lines, and this has been increasingly recognised as a customer satisfaction differentiator across all market segments, including luxury vehicles (Stylidis et al, 2016). This is confirmed by published results of global customer satisfaction surveys such as the JD POWER (2016) which show that in particular for the premium brands the quality of workmanship (fit & finish) play a significant role in brand choice. This confirms earlier research, such as Dagman et al (2004), which concluded that customers take high vehicle quality assessment of split lines for granted when buying premium cars.

This work is particularly focussed on customer perception of exterior craftsmanship quality in particular in relation to alignment of body parts. "Gaps", also referred to as "split lines", are commonly defined as the "relation between two mating parts over a specified distance” (Dagman et al, 2004). Figure 1 provides a graphical illustration of the alignment between the tailgate and the bodyside of a vehicle, which identifies the gap as a measurable quality characteristic. Given the geometrical complexity of the automotive body structures, designed to meet both aesthetics as well as functional requirements such as safety and aerodynamics, the achieved gap is the result of complex tolerance chains. Several recent studies have discussed the relationship between tolerancing and dimensional variability in split lines and perceived quality (Stylidis et al, 2017; Hoffenson et al, 2015).

From an automotive product development point of view, setting up targets for split lines early in product design is desirable, as this will enable a consistent approach to geometric design and tolerancing, as well as tool development to ensure that the gap targets are robustly achieved. However, insight of customer preference for split lines for specific body styles, in particular for new products, is often not available early in the vehicle programme, and targets are set based on designers' experience. This often leads to changes in specifications later in the design and development process, when the design can be evaluated against the known best in class based on prototypes. Late changes to targets are known to lead to product and programme quality issues, resulting in late delivery of engineering, additional costs, or late trade-offs to other targets causing their own downstream effects (Harris, 2016). Therefore, it is desirable to acquire a robust understanding of customer preference in relation to split lines, such that targets can be set "right first time" early in the vehicle programme.

The research presented in this paper was carried out from the perspective of an automotive OEM and was aimed at establishing and validating an efficient and effective methodology that can be deployed within a product development organisation, to support the robust capture of customer perception regarding gaps/split lines of car body panels. The approach used for the research was to use a case study, relating to a particular vehicle, for which specific questions were formulated, aiming to support the design specification for gaps / split lines in terms of nominal and tolerance range, as follows:

1. Do customers have a preference for smaller gaps? If so, at what gap value does customer appeal starts to diminish?
2. What variation in gap value becomes noticeable?

The organisation of the paper is as follows: section 2 presents an overview of related literature and research; section 3 summarises the case study methodology; section 4 presents the results and analysis from the study, followed by a discussion and summary of conclusions and recommendations.

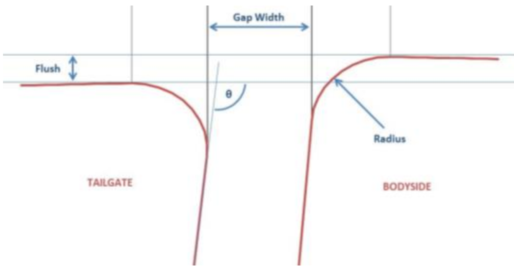



**Figure 1. Definition of Gap or Split Line as a measurable characteristic**

## 2. Critical analysis of related literature

Engineering perceived quality into design products has been significantly shaped by an emerging body of research aimed at establishing a comprehensive understanding of quality (Bjorn et al, 2017; Stylidis et al 2015) and methodological developments to enable systematic study of perceived quality for product development. This section provides a critical review of published methodology relevant to this study.

Kansei or Affective engineering aims to establish parametric links between the customers' emotional responses and the design characteristics of products or services (Nagamachi, 2008). Affective engineering studies are often based on focus groups (see for example Henson et al, 2006), where open ended questions are used to define the semantic space of words / adjectives that describe the characteristics of the product, followed by successive ranking and direct choice to select the most representative ones. The semantic differential method has also been illustrated by Hsu et al (2000) in a study to differentiate perceptions between designers and customers. The practical use of affective engineering in industry is limited both by the fact that it is resource intensive, and has limitations in relation to its use in complex systems where multiple feelings and quality dimensions interact with each other, which renders the affective product design research too complicated (Ayas, 2011).

Discrete choice experiments have been introduced as an efficient approach to study user preference for multi-dimensional product attribute context. The strength of this approach stems from the use of factorial Design of Experiments methodology to generate the set of choices for which customer preference choices are elicited, allowing for a better statistical evaluation of effects (Burgess & Street, 2005). Refined versions of discrete choice surveys have been increasingly used in perceived quality studies; e.g. Hoffenson et al (2015) have discussed the use of a choice based conjoint survey technique to study the customer preference in relation to a mix of attributes including build quality (in relation to geometric tolerances), price and sustainability in a global study. Johnson et al (2013) have presented a comprehensive review and recommendations for good practice for the experimental design of discrete-choice experiments, with comprehensive examples from health.

The strength of discrete choice experiments approach is that it is generally more efficient than affective engineering as it is less cognitively demanding for the participant. However, a limitation of discrete choice experiments is that it only collects information on a preference from many choices (two or more), which in turn might limit the practical relevance of the study. Increasing the choice set through contingent ranking which requires all choices to be ranked can be more cognitively demanding for the participant which in turn will impact on the effectiveness of the study.

Dagman et al (2004) investigated if there is a correlation between the car brands that customers rated high on aesthetics regarding split lines and actual measurement data of the geometrical variation of split lines of those car brands. Such correlation studies are very powerful as they allow for the study of more variable data, which results in better resolution of information and inference. However, this approach requires extensive data, which is both expensive due to the number of measurement data required, and difficult to obtain early in the design process for a specific product. Several researchers have discussed the use of combination of techniques including both attribute assessment and discrete choice experiments. For example Forslund et al (2013) presented a study focussed on identifying whether tolerance related geometrical deviations have an impact on customer preference. The study used photographs of the design artefacts considered (cordless telephones and bread toasters), and the survey systematically examined: (i) the attribute assessment of appeal over a 10 points rating scale; (ii) a preference test where participants were asked to choose between design pairs where the gap was the only difference, but without drawing participant's attention to the fact that the gap was the only difference; and (iii) an evaluation of preference after the participants were made aware that the gap was the only difference between the two designs.

Table 1 summarizes the studies discussed above on evaluating customer perception, detailing the survey working aids (types of display) and the type of survey, as well as the size of the study.

In relation to the use of visual graphics of products for research study, Artacho-Ramirez et al (2008) concluded that photographic representation suffices to communicate most of the concepts in the same way that the real product would do. While virtual prototypes are increasingly used by organisations to understand the user's needs during the initial phase of product development (Da Silva et al, 2016), in terms of testing the perception of visual quality 2D images provide sufficient visual information, and 3D images can only be justified if perception of stereoscopic depth is required (Ware, 2013).

Ware (2013) also discussed that the useful field of view (UFOV) is an important consideration for experiments to enhance visual information for participants from static images. UFOV is larger with low target density (i.e. less information displayed) and this means that the participant can take information quickly from larger area. Bergstrom et al (2016) discussed research results showing that older adults looked less frequently at peripheral parts of the screen, and therefore it is recommended to place the target in the central area of the screen. Nagy et al (1992) discussed that when the luminance or colour difference between target and distractors was high (i.e. when the target was more discriminable), it resulted in efficient search (larger UFOV) for the participants. It was also discussed that the UFOV can be achieved by the use of less saturated colours for large areas of display and highly saturated colours for small areas of display, to maximise discrimination.

Kantowitz et al (2014) discussed that demand characteristic can induce a Hawthorne effect (i.e. behaviour modification in response to awareness of being observed), and result in a significant uncertainty in relation to the results of the experiment. Researchers (e.g. Fostervold et al, 2001) have therefore introduced designed experiments to intentionally test for any demand characteristic effect.

Several researchers (see for example Dretske, 2006; and Merikle et al, 2001) have discussed the evidence for implicit, subliminal perception without awareness of the attributes of the object studied. Therefore, it is important to study perception without awareness to understand if participants are able to perceive a stimuli (such as the gap to body panel for a vehicle) without awareness, and to evaluate the implication that awareness might bring in relation to the choices made by survey participants.

**Table 1. Critical summary of studies on customer perception of product attributes**

| Reference | Method / Aim of study | Type of display | Mode of survey | Type of questions | Sample size |
|---|---|---|---|---|---|
| Henson et al (2006) | Affective engineering: Subjective requirements of packaging - shapes, volumes, textures & colours | CAD image - Grey colour | Semantic differential questionnaire; Video recording | Ranking; Open ended; Discrete choice | Focus group |
| Hsu et al (2000) | Affective engineering: Differences in product form perception between designers and users | Actual products (phones) | Semantic differential method | Ranked / 9 point scale; Open ended questions | 40 - 20 designers, 20 users |
| Hoffenson et al (2015) | Choice-based conjoint survey: Relationship between built quality and other attributes (price, sustainability) | Images - screen (mobile phones) | Online survey - conjoint experiment | Open questions + Discrete choice | Large scale - > 250 in 3 countries |
| Forslund et al (2013) | Combination of Discrete Choice + Attribute Assessment: Consumer assessment of product attributes based on visual quality appearance; | Photographs of products (toasters & phones) | Web based questionnaire + interview | Open ended; Ranking (10 points scale); Discrete choice | 48 participants (50% female / 50% male) |
| Dagman et al (2004) | Correlation study: Customer perception and branding impact of visual quality appearance in relation to split lines | Vehicle (actual) | Motor-show questionnaire | Open & closed ended questions | 130 participants |

# 3. Methodology

## 3.1. Case study, research objectives and approach to survey

The specific case study considered for this research is a sport utility vehicle. Table 2 below provides an exemplary illustration of the possible changes in design targets set for two different locations along the same vehicle split line, at different stages in the product development (i.e. an initial design target, and a revised target after the preliminary design was completed and dimensional variation analysis (DVA) conducted to evaluate the impact of manufacturing tolerances). With reference to this specific vehicle application the study aims to establish:

1. Whether customers have a preference for smaller gaps and whether there is a threshold for the gap value beyond which the customer appeal starts to diminish; this would support the specification of a nominal value for the gap; and
2. What variation in gap becomes noticeable to the customers - such that appropriate specifications for tolerances can be set?

Given the nature of the problem, and the desire to have an efficient process compatible with an OEM product development process / vehicle programme, a survey based on a combination of a discrete choice experiment and direct attribute assessment via ranking, similar to Forslund et al (2013), was chosen as the base methodology. The review of literature strongly suggested that several factors should be taken into account in developing robust discrete choice experiments, including:

3. Awareness of gap: do customers perceive gap without awareness? And does awareness of gap influences their preference for smaller or larger gap specification?
4. The impact of Hawthorne effect: does demand characteristic have an effect on the behaviour of respondents during the survey, leading to significant changes?

The design of the survey experiments needs to take into account the requirements / objectives outlined above. The process used by Forslund et al (2013) has addressed similar issues, and therefore was adopted as a blueprint for the current study. Specifically, the study of customer preference for gap was planned as a three phases survey, as followed:

- ***Phase 1***: Assessment of aesthetic appeal of gap size.
- ***Phase 2***: Discrete choice test aimed at identifying perception of gap and design preference without awareness.
- ***Phase 3***: Elicitation of preference with awareness of gap as target of evaluation.

The design of the experiment test cases should also allow for testing for any Hawthorne effect.
In the first instance, survey participants will be selected from employees in the company, to facilitate a faster turn-around of the study, while the methodology is developed and validated. However, the selection of participants will aim to ensure an appropriate gender and background (i.e. engineering vs non-engineering) balance to objectivise the study as much as possible.
Selecting an appropriate size of the survey in terms of the number of participants is an important study design consideration to ensure a good balance between statistical efficiency and response efficiency (Johnson et al, 2013). Testing for significance of the difference in mean ratings between options can be based on the principle of a 2-sample t-test, providing a sample size requirement for the experiments. Considering a power of the test of 0.70 in relation to a shift of 1 standard deviation, and a reference value for standard deviation of 1.97 based on the large scale research of vehicle styling and aesthetics appearance of Srinivasan et al (2012), leads to a required sample size of 49. This is compatible with the resource available to commit for this study, and was therefore adopted as target sample size.

**Table 2. Example of gap design specifications**

| Requirements | Initial Target | Revised Target (post DVA) |
|---|---|---|
| Location 1 | 3.5 ± 1 [mm] | 4.6 ± 1.6 [mm] |
| Location 2 | 3.5 ± 1 [mm] | 4.6 ± 1 [mm] |

## 3.2. Choice of Visual Graphics

Conducting a customer perception study based on vehicle physical prototypes would be both prohibitively expensive and infeasible early at the design stage. Visual graphics have been shown to provide a viable alternative, and as discussed by Ware (2013), static images of virtual prototypes adequately communicate concept in the same way in which the real product would do. Therefore, for the purpose of this study, AVA parametric CAD modelling virtual prototyping tool was used to generate images of the product showing the area of interest (tailgate to body side panel assembly). Figure 1 illustrates a partial image of the virtual prototype used in the study, exemplifying the way the UFOV principles have been used in creating the test samples.

## 3.3. Test cases and survey plan

Using the AVA parametric CAD modelling tool, images of the body assembly corresponding to different gaps between the tailgate and body panel were created. The values for gaps were chosen to represent the range of the tolerances illustrated in Table 2; five images were created, with gap values (in mm) as follows: 2.5; 3.0; 3.5; 4.6; 6.2.
For the preference test, where image pairs are shown to survey participants, a number of 8 test cases (labelled C1 to C8) were set up, summarised in Figure 3. The experimental set up tests 6 combinations of gaps contrasting nominal vs min / max tolerance and min vs max tolerances, including the extreme case of min/max tolerance across the 2 sets of targets (C8). Case C1 (pairing an image illustrating a 3.5mm gap with an image illustrating a 4.6 mm gap - corresponding to the nominal design values for the gap shown in Table 2 for different design stages) was repeated as case C3 in order to test for the Hawthorne effect.
Based on the 8 cases, the survey study was conducted in 3 phases:

1. ***Phase 1***: Assessment of aesthetic appeal of gap size: the images with different gap values were shown to the participants, who were asked to rate their assessment of "Aesthetics / Appeal" of each design on a scale of 1-10 (where 1 denotes is "Highly Unappealing", and 10 "Highly Appealing"). The images were arranged according to the 8 test cases; thus each

participant was asked to rate 16 images, each image being repeated at least once. The order of the cases was randomised to ensure repeat assessments of same image were not consecutive.

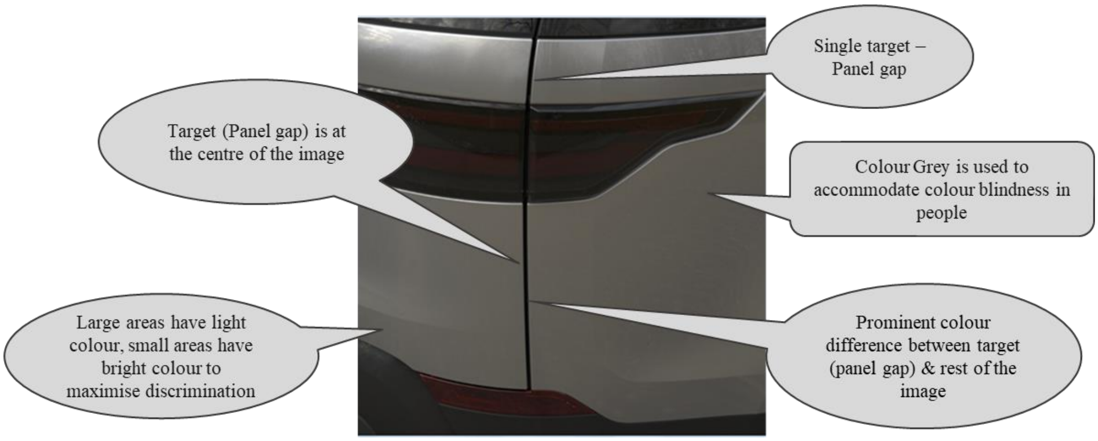



**Figure 2. Visual graphics used for the study, illustrating UFOV concepts employed**

2. ***Phase 2***: Discrete choice and preference test - without awareness: participants were shown two images next to each other (corresponding to the cases C1 to C8 in Figure 3), and were asked firstly to comment whether they note a difference between the 2 cases, without bringing the gap to the participants' attention. Respondent were also asked to indicate their preference for one image / design versus the other - again without any reference to the gap.
3. ***Phase 3***: Preference with awareness: in this phase the participants were shown again the test cases image pairs and were made aware that the gap was the only difference between the two otherwise identical designs, and asked to indicate again their preference for one image / design versus the other.

| **Gap** [mm] | **2.5** | **3.0** | **3.5** | **4.6** | **6.2** |
|---|---|---|---|---|---|
| **2.5** | | | **C2** | **C5** | **C8** |
| **3.0** | | | | **C4** | **C7** |
| **3.5** | **C2** | | | **C1,3** | |
| **4.6** | **C5** | **C4** | **C1,3** | | **C6** |
| **6.2** | **C8** | **C7** | | **C6** | |

**Figure 3. Test cases planned for the study (Gap size in mm)**

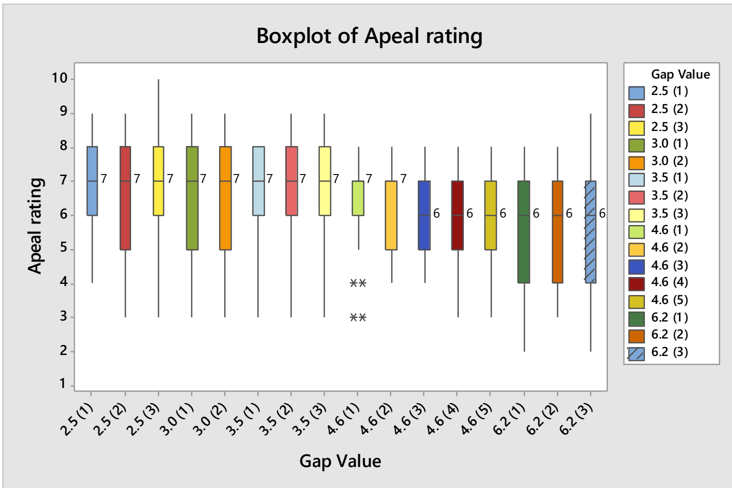
## 4. Results and analysis

In the actual study, deployed in 3 phases as discussed above, full results were only available from 47 participants. A good gender balance was maintained, with 23 female and 24 male participants. The survey participants were employees of the Company, selected from all departments (including non-engineering departments such as administration, purchasing, corporate strategy), ensuring balance between engineering and non-engineering backgrounds of the participants. The following subsections summarise the analysis of the results against the research questions set for the study.

## 4.1. Influence of gap size on aesthetic appeal

Figure 4 summarizes the statistical analysis of the result for the evaluation of impact of gap size on customer preference, collected as rating for the aesthetic appeal in Phase 1 of the study.
The statistical summary of the results shown in Figure 4.a show a consistent standard deviation of ratings across the experiments, and also consistent with evaluations from much larger studies discussed in literature (Srinivasan et al, 2012). Taking a worst case standard deviation of 1.8 (from Figure 4.a), with the actual sample size of 47 participants, points to an actual power of a 2 sample t-test of 0.76 against a difference of 1 standard deviation, which is acceptable against the assumptions set for the study.
Figure 4.b presents a boxplot analysis of the appeal ratings provided by the participants across Phase 1 of the study. Each of the gap values were rated more than once by the participants (e.g. gap value 2.5 was rated 3 times, with results denoted as 2.5(1), 2.5(2), and 2.5(3) in the labels of the x axis in Figure 4.b). While the results in Figure 4.b are presented in the ascending order of the gap, the participants were shown the images in the order of gaps values studied in each of the 8 test cases.
The results in Figure 4.b show that the participants as a group seem to be consistent in their appeal rating for each of the gap values; this is shown by the median appeal score and interquartile range across the repeated evaluations of the same gap. It is also noticeable from Figure 4.b that the appeal appears to diminish at gap values of 4.6 and larger.

| Variable | Mean | StDev | Median |
|---|---|---|---|
| Gap 2.5 | 6.532 | 1.487 | 7.000 |
| Gap 4.6 | 5.957 | 1.429 | 6.000 |
| Gap 2.5_1 | 6.766 | 1.577 | 7.000 |
| Gap 6.2 | 5.617 | 1.800 | 6.000 |
| Gap 3.0 | 6.447 | 1.530 | 7.000 |
| Gap 4.6_1 | 6.074 | 1.327 | 6.000 |
| Gap 3.0_1 | 6.574 | 1.543 | 7.000 |
| Gap 6.2_1 | 5.553 | 1.730 | 6.000 |
| Gap 3.5 | 6.511 | 1.428 | 7.000 |
| Gap 2.5_2 | 6.798 | 1.401 | 7.000 |
| Gap 3.5_1 | 6.511 | 1.381 | 7.000 |
| Gap 4.6_2 | 6.468 | 1.283 | 7.000 |
| Gap 3.5_2 | 6.511 | 1.397 | 7.000 |
| Gap 4.6_3 | 6.266 | 1.310 | 7.000 |
| Gap 4.6_4 | 6.053 | 1.475 | 6.000 |
| Gap 6.2_2 | 5.723 | 1.690 | 6.000 |




a. Statistical summary of results b. Boxplot analysis of gap appeal ratings

**Figure 4. Results analysis: Appeal rating for gap**

A further study was conducted based on the Phase 1 test results to evaluate the change in gap value that leads to a significant change in the appeal rating. A non-parametric Mann-Whitney U-test, commonly used with Likert data (Forslund et al, 2013), was conducted to test the significance of the difference between case pairs of gap size appeal ratings. Figure 5 summarizes the achieved significance values for the Mann-Whitney test for each pairwise comparison of gap appeal rating for cases C1 - C8. These results indicate that gap differences below 2 mm do not appear to significantly influence the appeal rating. For case C5 (2.5mm vs 4.6mm - roughly corresponding to the initial design tolerance range - see Table 2), at a gap difference of 2.1 the significance is 0.057, which could be taken as significant - and hence conclude that 2.1 mm difference in gap will result in a significant appeal rating difference. Quite clearly gap differences bigger than 3mm appear to be related to significant differences in appeal rating.

## 4.2. Study on the influence of the Hawthorne effect

In Phase 2 of the study where the participants are specifically asked to look for differences between images and to express a preference for one design choice versus another, it is likely that participants may feel to be under focus, especially as they are asked this question in total 8 times corresponding to the 8 test cases C1 - C8. In this situation any influence of demand characteristic / Hawthorne effect is important to be evaluated in order to prevent making any exaggerated or incorrect inference. In order to facilitate this study, a repeat of case C1 was introduced as case C3, i.e. the participants were presented with the same pair of images as in C1. Any significant difference in results will indicate a potential significant Hawthorne effect on participant response. Figure 6 summarises the results of the study in terms of the percentage of respondents that have indicated the gap as the difference between the 2 images shown.

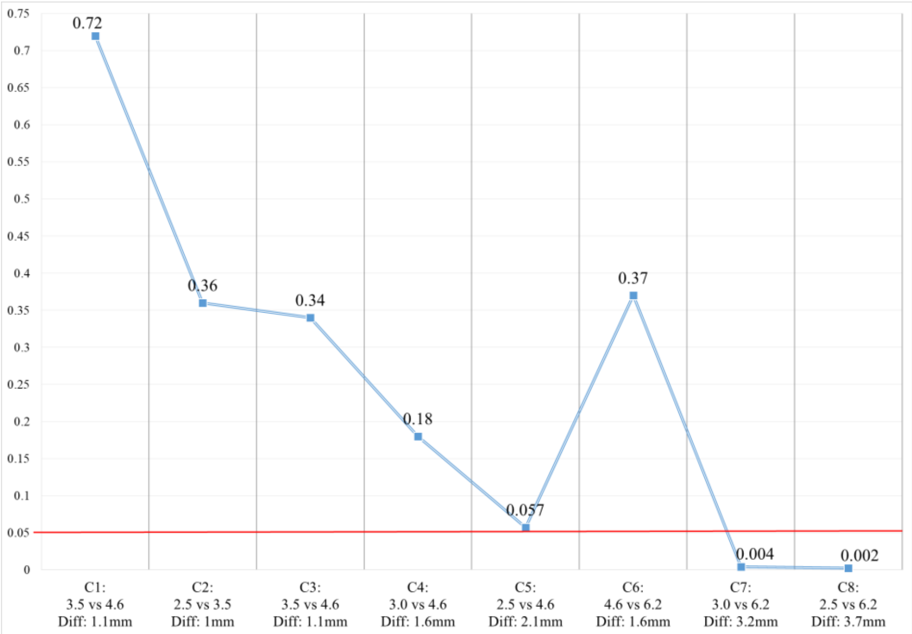



**Figure 5. Significance values for Mann-Whitney U tests for C1-8 gap pairs**

The results in Figure 6 are significant in two ways: firstly, there is a marked difference between the results for the repeat test case of the gap pair 3.5mm vs 4.6mm (corresponding to the 2 nominal gap targets in Table 2). In test case C3, 31.91% of the participants noticed gap as a difference, whereas in the test case C1 (which was the first test in the series) only 12.77% of the participants noticed gap as a difference between the 2 prototype images.

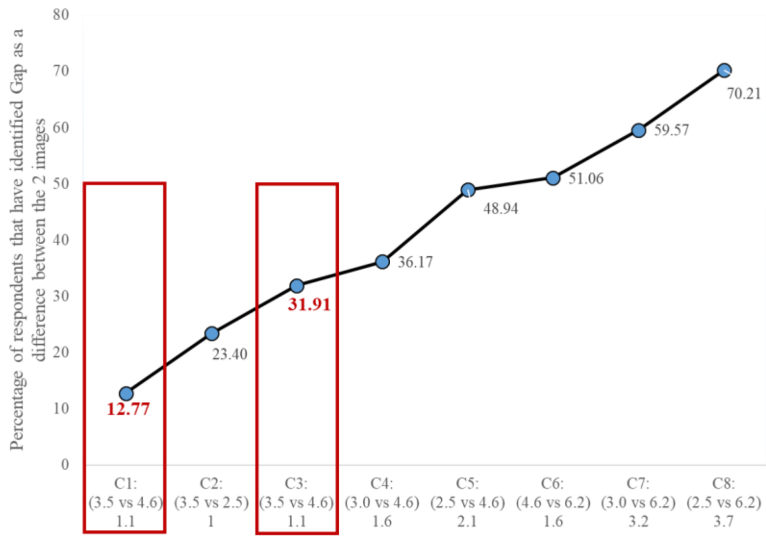



**Figure 6. Influence of Hawthorne effect in the study**

This suggests that the Hawthorne effect has an influence on the participants' response. Figure 6 also indicates that the proportion of participants noticing the difference increases with the progression through the test cases in the study, which is arguably only partly explained by the gap difference. As well as the difference between C1 and C3, it is also seen that there is a difference between C4 and C6 where a similar gap difference is perceived by a higher proportion of respondents.

### 4.3. Perception without awareness of gap size

During Phase 2 of the study, participants were asked whether they observe a difference between the 2 images of rear vehicle body and asked to express a preference between the 2 images of design. For the participants that did not identify gap as a difference between the 2 designs, the analysis of their preference for smaller or larger gap designs is important as it reveals the gap perception without awareness. Figure 7 summarises the results from the study in a matrix format.

The results show that in most cases (4 out of 7 distinct cases) the participants have indicated a preference for the smaller gap image / design without awareness. However, the weight of preference is quite small except for cases C2 and C6. Quite interestingly, for the larger gap differences - cases C5, C7 and C8, the weight of the preference was split between small gap and large gap - without awareness.

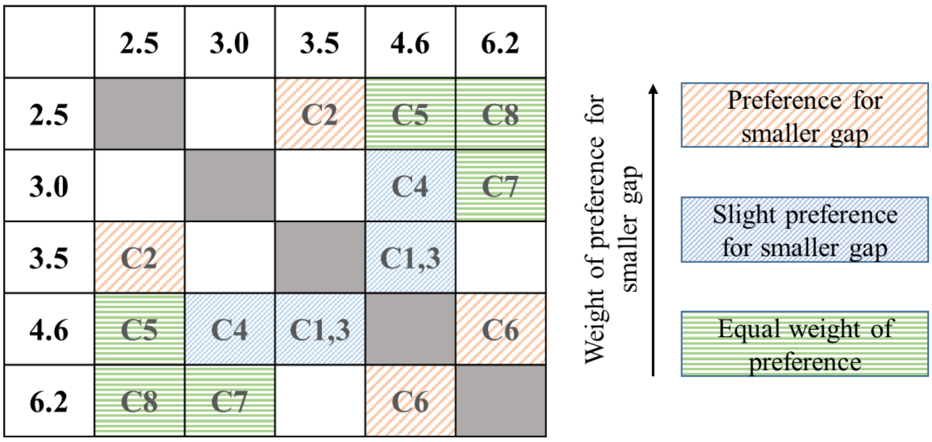



**Figure 7. Results of gap size perception without awareness**

### 4.4. Shift in gap preference with awareness

In phase 3 of the study, the participants were made aware of the fact that gap is the only difference between the 2 images / designs they were asked to evaluate. This enabled a study in the change in behaviour of the respondents with awareness of the gap size, in particular to evaluate whether there is a clear preference for smaller gap from the participants. While some participants have not changed their preference after they have made aware of the gap difference, many have re-evaluated their preference. Figure 8 summarizes the analysis of the change in proportion of preference for small gap "before" and "after" the participants were made aware of the gap difference for each of the test cases C1 - C8, by means of a multi-vari chart.

The results in Figure 8 indicate a clear shift in preference for small gap after the participants have been made aware of the gap being the only difference. There is an overall mean shift in the percentage of participants that expressed preference for small gap across all test cases, from 44% "before" to 82% "after" the discussion of gap. It is also noted that the shift of preference for small gap is evident in all test cases; the lowest proportion of 69% was recorded in test case C2, which had the smallest gap difference (2.5mm - 3.5mm). For larger differences in gap and for larger gaps (test cases C6 - C8) the proportion of preference for small gap is larger than 90%.

## 5. Discussion, Conclusions and Recommendation

The research presented in this paper was carried out from the perspective of an automotive OEM, and a main objective was to establish an efficient and effective methodology for evaluation of perceived quality in relation to body split lines / gaps, early in the product development process, such that targets and tolerances can be appropriately set for a target vehicle. The methodology for the case study that underpinned the research was carefully established based on the review of best practices available through the academic literature, and consisted of a combination of direct attribute evaluation and choice experiments with multiple test cases, based on static images generated from parametric CAD virtual prototypes. Based on the results from the study, the key reflections and conclusions in relation to the methodology are as follows:

1. The use of static images from parametric CAD models is suitable for evaluation of perceived vehicle body quality. This is important because parametric CAD models can be used to generate a range of virtual prototypes that can be evaluated early in the product development process for a specific vehicle.

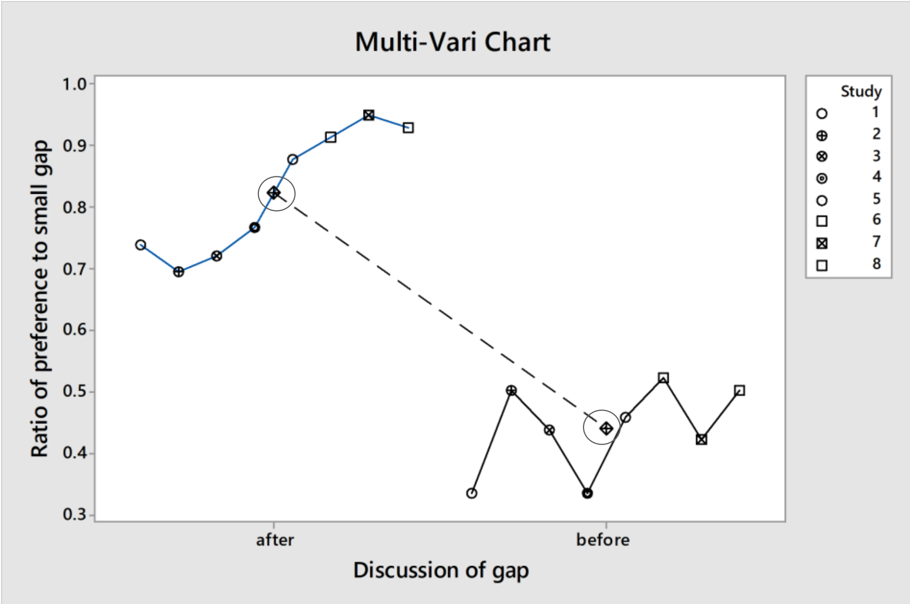



**Figure 8. Multi-Vari chart of change in preference for small gap**

2. The approach of using Useful Field of View concept to enhance information visualisation for participants from static images of parametric CAD model has enabled to eliminate any negative effect of poor visualization and has thereby helped to ensure better cognitive response from participants.
3. The study showed that the combination of direct attribute rating and choice experiments is effective and necessary in comprehensive capture of user preference and evaluation of perceived quality. The analysis of aesthetic appeal rating data has provided an indication that the appeal rating worsens with larger gaps. The pairwise statistical significance study was useful in identifying whether variation of gap within tolerance limits (>2.1mm for the specific case study considered in this paper) could lead to a significant change in appeal rating. The choice experiments have provided clear evidence of the user preference for smaller gap size, in particular when the participants were made aware that the gap is the only difference.
4. The study clearly showed the importance of the planning of the experiments in relation to the condition of the participants, and the impact this can have in relation to the conclusions from the study. This relates to both the demand characteristic / Hawthorne effect, which was shown that can have a significant impact, and the effect of attribute awareness in relation to stated preference. The study on perception without awareness within this experiment on 8 test cases has shown that the stimuli gap to body panels is not perceived without awareness.
5. For the particular perceived quality gap problem studied in this paper, given the significant shift in participant preference once the awareness of gap as the only difference is made, makes it a clear conclusion that the size of the gap is an important "the smaller the better" type quality characteristic. This is coherent with the findings from the appeal rating - and substantiates gap as the differentiator characteristic. Given that the Hawthorne effect was tested across the 3 phases of the study (by repeating a test case without the participants to be

noted about this), this conclusion cannot be put under doubt by any demand characteristic effect; indeed, Figure 8 clearly shows that while there is a difference in preference between cases C1 and C3 (i.e. the repeated test case), the size of the shift in preference after gap awareness was made, makes it clear that this is not due to a demand characteristic effect. It is interesting to note that in the Phase 3 experiments there is virtually no difference between the results from C1 and C3 - which means that with awareness there is no Hawthorne effect (presumably due to the greater confidence of the participants in expressing their preference against a more clearly defined context for the choice).

6. The approach taken for the planning of the experiments - by carefully defining test cases representing realistic contrasts of gap values that can be seen given the hypothesised tolerance specifications, and allowing repeat test cases, has been proven to be good practice. This has allowed for repeat evaluations of the same image or image pair - which allowed a better estimation of the effects from the experiment, including accounting for any Hawthorne effect.

The above points clearly define the contribution of this study to the academic body of knowledge and practice guides in relation to perceived quality experiments for automotive body design.

The methodology can be replicated for other similar or related problems in relation to vehicle body perceived quality, e.g. to test for user preference with split lines / gap for different body styles, or to evaluate appeal and preference in relation to other attributes such as flush, radius, draft angle, etc. The experiment can be (and has been) replicated with other groups of participants to evaluate any difference in appeal and preference between customer groups.

## 6. References


Artacho-Ramirez, M.A., Diego-Mas, J.A. and Alcaide-Marzal, J., (2008) Influence of the mode of graphical representation on the perception of product aesthetic and emotional features: An exploratory study. International Journal of Industrial Ergonomics, 38(11), pp.942-952, DOI: 10.1016/j.ergon.2008.02.020.

Ayas, E., (2011) Engineering feelings of quality: Applications for products, service environments and work systems, Doctoral thesis, KTH Stockholm.

Falk, B., Stylidis, K., Wickman, C., Soderberg, R., Schmitt, R., 2017, Shifting paradigm: towards a compreheensive understanding of quality, Proc. 21st Int. Conf on Eng Design, ICED17, Vol.9 255-264.

Burgess, L, Street, D.J, (2005) Optimal designs for choice experiments with asymmetric attributes, Journal of Statistical Planning and Inference 134 (2005) 288–301, DOI: 10.1016/j.jspi.2004.03.021.

Dagman, A., Wickman, C. and Söderberg, R., (2004), September. A Study of Customers' and the Automotive Industry's Attitude Regarding Visual Quality Appearance of Split-Lines. In 4th International Conference on Advanced Engineering Design AED.

Da Silva, G.C. and Kaminski, P.C., (2016). Selection of virtual and physical prototypes in the product development process. The International Journal of Advanced Manufacturing Technology, 84(5-8), pp.1513-1530.

Dretske, F., (2006) Perception without awareness, chapter in "Perceptual Experience", Eds Gendler, T.S. & Hawthorne, J., Oxford University Press, DOI: 10.1093/acprof:oso/9780199289769.003.0005.

Forslund, K., Karlsson, M. and Söderberg, R., (2013) Impacts of geometrical manufacturing quality on the visual product experience. International Journal of Design, 7(1), 69:84.

Harris, S (2016) Functional Design – Dimensional Management Compatibility tools mobility. MSc Dissertation, University of Bradford, UK.

Henson, B., Barnes, C., Livesey, R., Childs, T. and Ewart, K., (2006) Affective consumer requirements: A case study of moisturizer packaging. Concurrent Engineering, 14(3), pp.187-196, DOI: 10.1177/1063293x06068358.

Hoffenson, S., Dagman, A. and Söderberg, R., (2013) Tolerance optimization of a mobile phone considering consumer demand for quality and sustainability in china, sweden, and the united states. In Proceedings of the 19th International Conference on Engineering Design.

Hoffenson, S., Dagman, A. and Söderberg, R., (2015) Visual quality and sustainability considerations in tolerance optimization: A market-based approach, Int. J. Production Economics 168 (2015) 167–180, DOI: 10.1016/j.ijpe.2015.06.023.
Hsu, S. H., Chuang, M. C., & Chang, C. C. (2000). A semantic differential study of designers' and users' product form perception. International Journal of Industrial Ergonomics, 25(4), pp.375–391, DOI: 10.1016/s0169-8141(99)00026-8.
J D POWER and Associates (2016), U.S. Vehicle Dependability Study.
Johnson, F.R, Lancsar, E., Marshall, D., Kilambi, V., Mulbacher, A, Regier, D.A., Bresnahan, B.W., Kanninen, B., Bridges, J.F.P. (2013) Constructing experimental Design for Discrete Choice Experiments: Report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force, Value in Health, 16 (2013) 3-13, DOI: 10.1016/j.jval.2012.08.2223.
Kantowitz, B.H., Roediger III, H.L. and Elmes, D.G., (2014) Experimental psychology. Nelson Education.
Merikle, P.M., Smilek, D. and Eastwood, J.D., (2001) Perception without awareness: Perspectives from cognitive psychology. Cognition, 79(1), pp.115-134, DOI: 10.1016/s0010-0277(00)00126-8.
Nagamachi, M., (2008) Perspectives and the new trend of Kansei/affective engineering. The TQM Journal, 20(4), pp.290-298, DOI: 10.1108/17542730810881285.
Romano Bergstrom, J. C., Olmsted-Hawala, E. L., & Bergstrom, H. C. (2016). Older adults fail to see the periphery in a web site task. Universal Access in the Information Society, 15(2), 261-270, DOI: 10.1007/s10209-014-0382-z.
Srinivasan, R., Lilien, G.L., Rangaswamy, A., Pingitore, G.M. and Seldin, D., (2012) The total product design concept and an application to the auto market. Journal of Product Innovation Management, 29(S1), pp.3-20, DOI: 10.1111/j.1540-5885.2012.00958.x.
Stylidis, K., Landahl, J., Wickman, C., Johanesson, H, Soderberg, R., (2015) Structuring perceived quality attributes for use in the design process, Proc. ICED 15.
Stylidis, K., Madrid, J., Wickman, C., Soderberg, R., 2017, Towards overcoming the boundaries between manufacturing and perceive quality: an example of automotive industry, Procedia CIRP 63 (2017) 733-738, DOI: 10.1016/j.procir.2017.03.091.
Ware, C., (2013). Information visualization: Perception for design, 3rd ed, Elsevier Morgan Kaufmann.
Wickman, C., Wagersten, O., Forslund, K., Soderberg, R., 2014, Influence of rigid and non-rigid variation simulations when assessing perceived quality of split lines, Journal of Engineering Design, Vol 25: 1-3: 1-24, DOI: 10.1080/09544828.2013.875522.